DOCUMENT RESUME

ED 092 572 TM 003 674

AUTHOR Lieberman, Marcus
TITLE Psychometric Analysis of Developmental Stage Data.
PUB DATE [Aug 73]
NOTE 22p.; Paper presented at the Annual Meeting of the American Psychological Association (Montreal, Canada, August 1973)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Abstraction Levels; Behavior Theories; Complexity Level; *Developmental Psychology; *Measurement Techniques; *Moral Values; Psychological Patterns; *Psychometrics; Task Analysis
IDENTIFIERS *Latent Trait Theory

ABSTRACT

When subjects are given open-ended stimulus situations allowing for responses at all possible stages of a developmental theory, the stage scores of individuals for each situation can be treated as scores on polychotomous items. Extensions of the concepts of difficulty and discriminating power from the dichotomous case to this ordinal category situation allow determination of these parameters. Assuming the trait being measured is distributed normally, stage difficulties or boundaries are deviates corresponding to cumulative proportions of subjects at each stage. Item discriminating powers are determined by the one factor loadings of the inter-item correlation matrix. Maximum-likelihood estimates of a subject's trait level can be determined from the parameters of each situation presented and that subject's answer pattern of stage scores. Data gathered in interviews to determine the level of moral development according to Lawrence Kohlberg's six-stage theory takes the above form. Results indicate that for different dilemmas, certain stages of reasoning are more easily attained than others, and when particular moral issues are analyzed across situations, an "entry" issue for various stages is indicated corresponding to Kohlberg's theory. (Author/MLP)

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED 092572

TM 003 674

Psychometric Analysis

of

Developmental Stage Data

Marcus Lieberman

Harvard University

Graduate School of Education

Paper presented at the Annual Meeting of the American Psychological Association, Montreal, Quebec, Canada August, 1973.

Psychometric Analysis of Developmental Stage Data

Marcus Lieberman

Harvard University

Graduate School of Education

Introduction

Traditionally, psychometricians have analyzed tests composed of items or tasks to which responses may be classified as "pass" or "fail." Most typically, these are attitude or achievement measures often multidimensional in nature. The past few years have seen efforts focused on a new probabilistic, "latent trait" model of subject and item parameters, rather than the "true-score" model. Lord (1952) discusses item and subject parameter estimation and Birnbaum (in Lord and Novick, 1968) considers a number of mathematical models which depend on the distribution of the trait in the population of subjects. It is this second model which will be applied to stage-type data.

Stage theorists postulate a series of structured wholes, forming an invariant sequence where passage from an inferior to a superior stage is equivalent to an integration. Often, "stage specific" tasks are presented to a subject and if almost all of a subset are passed and few of a more complex set, the subject is identified at a specific stage. Being in a transitional state is manifested by passing some tasks at both of two adjacent stages. The alternative to "stage specific" tasks is the presentation to the subject of an open-ended task, sentence

or situation, to which responses may be classified as representative of a particular stage or mixture of two adjacent stages. The pattern of responses across several tasks yields an overall assignment to stage or transition.

This open-ended task, sentence or situation may be considered to be an item with a range of scoring possibilities rather than the dichotomous pass-fail classification. Logical extensions of techniques employed to analyze tests of dichotomous items may be applied to this polychotomous case. First, task difficulties and discriminating powers are determined; then, given a subject's response pattern to a number of tasks, a value for his latent trait which maximizes the probability of producing that pattern is estimated.

In the following sections the techniques for estimating item and subject parameters are discussed. It may be useful at this point, however, to see how they are related in a dichotomously scored example. Consider Figure 1.

The "characteristic curves" of three items are displayed. Item C is the most difficult of the three because it takes much more of the trait to even have a finite probability of answering the item correctly. Item B is the most discriminating, for below a certain level of the trait one has zero probability of answering the item correctly and above that level one has near certainty of doing so. Finite probabilities of answering Items A and C correctly exist over a large range of the latent trait.

The Item Parameters

The estimate of difficulty in the dichotomously scored item, when the trait of the population answering the item is assumed normally distributed, is the normal deviate corresponding to the proportion answering the item correctly. In extending

this notion to the polychotomous situation, the difficulties, or boundaries, are the normal deviates corresponding to the cumulative proportions of subjects scored at each stage. The hierarchical nature of the stages allows us to state that a subject at a given stage has passed through all the previous stages.

In dichotomously scored items, discriminating power is estimated in a number of different ways. The point biserial item-total correlation is one common method, often approximated by taking the difference between the proportion of top scoring subjects (based on total score) answering an item correctly and the proportion of bottom scoring subjects also answering that item correctly. A second method suggests using the first factor loadings of the assumed unifactoral matrix of inter-item correlations. In the dichotomous case, these correlations are tetrachoric coefficients.

In the multi-category situation, one might be tempted to use product-moment correlation coefficients, but examination of the boundaries of typical polychotomous items shows them to be non-equidistant in nature, and simple arithmetic procedures are thus inadequate. The correct coefficient is the polychoric correlation.

Pearson and Pearson (1922) first discussed this coefficient but computing procedures were highly complex. A maximum likelihood estimation was proposed by Tallis (1962) and a simplified form when heuristic estimates of difficulties are accepted rather than estimated is derived by Lieberman (1969).

Items are taken in successive pairs and a correlation coefficient is computed for each pair. After coefficients are produced for all possible pairs of items, they are grouped in the familiar matrix form. Then a factor analysis using a

single factor (an assumption which must be tested) will produce a loading for each item which will serve as a measure of that item's discriminating power.

The Subject Parameter

Once the difficulties (or boundaries) and discriminating power for each item are determined, a subject's answer pattern of stage scores across a set of items will produce an estimate of that subject's latent trait. The method described below will produce, for a subject randomly chosen from a population whose trait may be assumed normally distributed, the same trait value for any subset of items. However, the more stories employed, the smaller will be the standard error of that value.

The probability of being scored at a particular stage for a particular item is a function of a subject's latent trait value, the discriminating power of the item and the boundaries about that stage. Since the items are assumed locally independent, the probability of an entire pattern is merely the product of the probabilities of achieving each item score. For various values of the latent trait, it is possible to calculate the probabilities of a given pattern occurring, and the value that maximizes this probability is the maximum likelihood estimate of the latent trait value of a subject exhibiting this pattern. The likelihood equations and complete technique are described in Samejima (1969).

An Application: The Development of Moral Judgment

Kohlberg (1969) has postulated six stages in the development of moral judgment. These ways of reasoning about moral dilemmas are briefly described in Table 1. Researchers present a set of situations or dilemmas to subjects and probe their

responses so as to get a sufficient amount of data to categorize the level of reasoning according to one of the six stages. Originally there was a single stage score for each dilemma but more recently an issue scoring system has been devised which allows a scorer to read a subject's references to a single moral issue (e.g., the value of life, the importance of rules, etc.) across all dilemmas. Then stage scores on various issues are recorded and an overall stage is calculated. In the early research a weighted average was computed across the story or issue scores to arrive at a global score.

The Stories

A sample dilemma is given in Figure 2 accompanied by possible probe questions an interviewer might ask in order to more precisely determine the level of reasoning the subject might use.

Table 2 gives the stage boundaries and discriminating power for each of five dilemmas. It may be noted that some stories are more difficult than others at certain stages. Figure 3 graphically illustrates these boundaries and the non-interval distances between the stages is quite evident.

A dramatic example of how using more information from the data than an arithmetic average requires is pictured in Table 3. Here, patterns across the five stories yielding the same arithmetic average give different estimates of the latent trait due to the varying difficulties and discriminating powers of the stories.

The polychotomous analogy to the item characteristic curve described earlier is pictured in Figure 3. Instead of the probability that a subject will pass an

item, the probability that a subject will cross the boundary between two adjacent stages is presented. Note that the low discriminating power produces curves which give finite probabilities of crossing each boundary for a wide range of the latent trait, moral judgment level.

The value of the latent trait, as it stands, has little conceptual meaning other than to place each subject in a normative position to the population, though this is the variable that would be used in any group comparison analyses. To relate the value of the latent trait to the stages, one finds the probability that a subject with given trait value would respond to a "pure" pattern, i.e., 11111 through 66666 across the stories. This procedure reveals the probability that a subject is at any one of the six stages. A graphical representation of this relationship is given in Figure 4.

The Issues

The various moral issues raised in the dilemmas are given in Figure 5. Issue parameters are listed in Table 4 and boundaries are illustrated in Figure 6. Note the highly increased discriminating powers reflecting the consistent thinking of subjects on a single moral issue in varying dilemmas. This increase in discriminating power also is revealed in the issue characteristic curves given in Figure 7 and the stage probabilities in Figure 8.

Conclusions

A great deal of information is provided by the graded response model. Story or task differences in difficulty and discriminating power may be estimated. A continuous, normally distributed trait level can be determined with its standard error for group comparisons. Apart from these data, the method provides a

description of the subject's probability of being in each stage, given his response pattern to the stories.

These probabilities can be used in a variety of ways. For example, it may be interesting to know which children in a given group are "nearly ready" to change to the next highest stage. This occurs when the probabilities of a subject being in adjacent stages are nearly equal. This information will be valuable to experimenters attempting to change the level of reasoning with an intervention program of discussions, readings, or experiences.

There are many developmental concepts which are being viewed as sequences of stages where each additional stage attained implies not only a substitution or addition but an integration with earlier stages. Some of these are ego development, physiological development, etc. In each, situations are discussed, tasks attempted or observations made. The researcher than rates the subject at a particular stage of development for each task, situation or observation. Using the technique described in this paper, situation, task or observation parameters may be estimated and used to determine subjects' latent trait levels for contrasts among groups of varying characteristics.

References

Kohlberg, L. Stage and Sequence: The Cognitive-Developmental Approach to Socialization, Handbook of Socialization Theory and Research, David A. Goslin (Ed.), Rand McNally, 1969.

Lieberman, M. Calculation of a Polychoric Correlation Coefficient. Paper presented at the Spring, 1969 meeting of the Psychometric Society, Princeton, New Jersey.

Lord, F. A Theory of Test Scores. Monograph No. 7, Psychometrika, 1952.

Lord, F. and Novick, M. Statistical Theories of Mental Test Scores, Addison Wesley, 1968.

Pearson, K. and Pearson, E. S. On Polychoric Coefficients of Correlation, Biometrika, 1922, 14, 127-156.

Samejima, F. Estimation of a Latent Ability Using a Response Pattern of Graded Scores. Monograph No. 17, Psychometrika, 1969, 34: 4(2).

Tallis, L. R. The Maximum Likelihood Estimate of Correlation from Contingency Tables, Biometrics, 1962, 18, 342-353.

Figure 1

Item Characteristic Curves

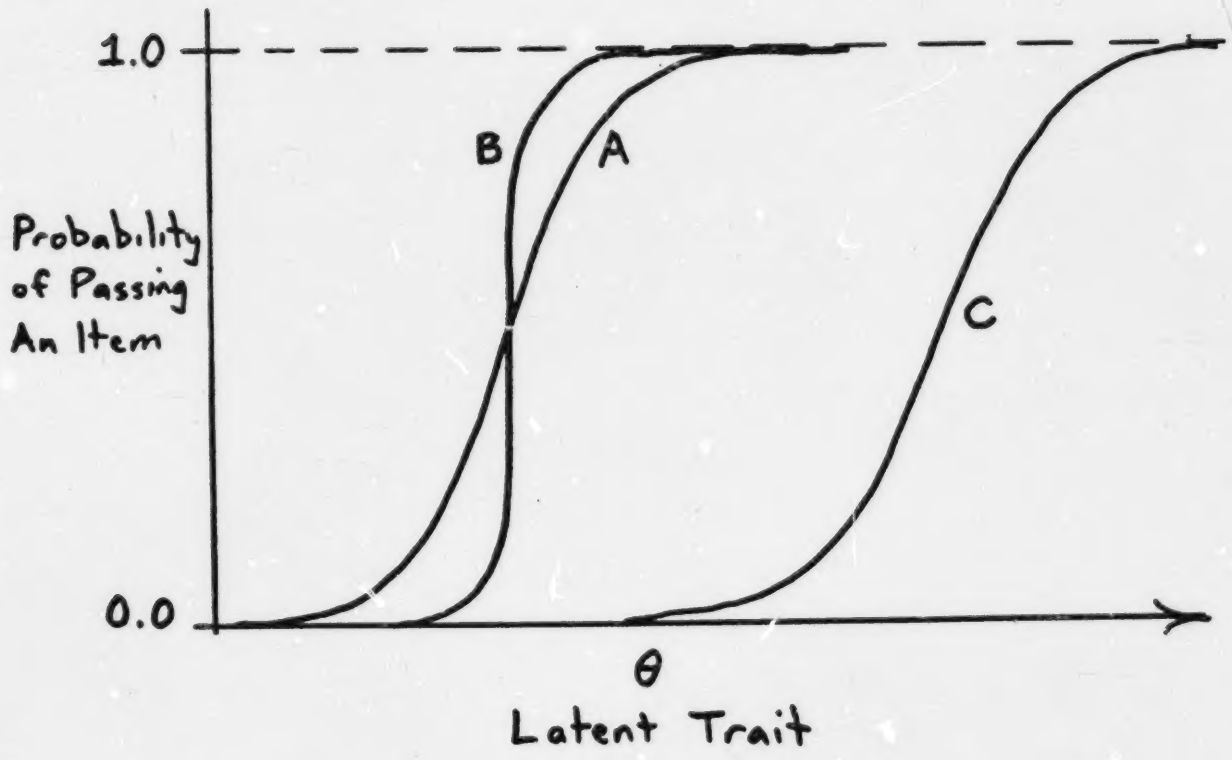


Table 1

CLASSIFICATION OF MORAL JUDGMENT INTO LEVELS AND STAGES OF DEVELOPMENT

Levels	Basis of Moral Judgment	Stages of Development
I	Moral value resides in external, quasi-physical happenings, in bad acts, or in quasi-physical needs rather than in persons and standards.	Stage 1: Obedience and punishment orientation. Egocentric deference to superior power or prestige, or a trouble-avoiding set. Objective responsibility.
		Stage 2: Naively egoistic orientation. Right action is that instrumentally satisfying the self's needs and occasionally others'. Awareness of relativism of value to each actor's needs and perspective. Naive egalitarianism and orientation to exchange and reciprocity.
II	Moral value resides in performing good or right roles, in maintaining the conventional order and the expectancies of others.	Stage 3: Good-boy orientation. Orientation to approval and to pleasing and helping others. Conformity to stereotypical images of majority or natural role behavior, and judgment by intentions.
		Stage 4: Authority and social-order maintaining orientation. Orientation to "doing duty" and to showing respect for authority and maintaining the given social order for its own sake. Regard for earned expectations of others.
III	Moral value resides in conformity by the self to shared or shareable standards, rights, or duties.	Stage 5: Contractual legalistic orientation. Recognition of an arbitrary element or starting point in rules or expectations for the sake of agreement. Duty defined in terms of contract, general avoidance of violation of the will or rights of others, and majority will and welfare.
		Stage 6: Conscience or principle orientation. Orientation not only to actually ordained social rules but to principles of choice involving appeal to logical universality and consistency. Orientation to conscience as a directing agent and to mutual respect and trust.

Figure 2

Story III.

In Europe, a woman was near death from a special kind of cancer. There was one drug that the doctors thought might save her. It was a form of radium that a druggist in the same town had recently discovered. The drug was expensive to make, but the druggist was charging ten times what the drug cost him to make. He paid $200 for the radium and charged $2,000 for a small dose of the drug. The sick woman's husband, Heinz, went to everyone he knew to borrow the money, but he could only get together about $1,000 which is half of what it cost. He told the druggist that his wife was dying, and asked him to sell it cheaper or let him pay later. But the druggist said, "No, I discovered the drug and I'm going to make money from it." So Heinz got desperate and broke into the man's store to steal the drug for his wife.

1. Should Heinz steal the drug? Why?

2. Which is worse, letting someone die or stealing? Why?

2a. What does the value of life mean to you, anyway?

3. Is there a good reason for a husband to steal if he doesn't love his wife?

4. Would it be as right to steal it for a stranger as his wife? Why?

5. Suppose he was stealing it for a pet he loved dearly. Would it be right to steal for the pet? Why?

6. Heinz steals the drug and is caught. Should the judge sentence him or should he let him go free? Why?

7. The judge thinks of letting him go free. What would be his reasons for doing so?

8. Thinking in terms of society, what would be the best reasons for the judge to give him some sentence?

9. Thinking in terms of society, what would be the best reasons for the judge to not give him some sentence?

TABLE 2

STORY PARAMETERS

Story		I	III	IV	VII	VIII
Disc Power		.3167	.4817	.3370	.4474	.5769
Stage Bound.	1	-1.16577	-1.15902	-1.34213	-1.15567	-1.57095
	2	-0.18472	-0.43518	-0.06652	-0.58667	-0.67087
	3	0.50850	0.39344	0.64134	0.49074	0.27614
	4	1.15297	1.35264	1.58417	1.28556	1.11006
	5	2.93346	3.33138	2.06890	2.60977	3.58486

Figure 3
Category Boundaries for Five Stories

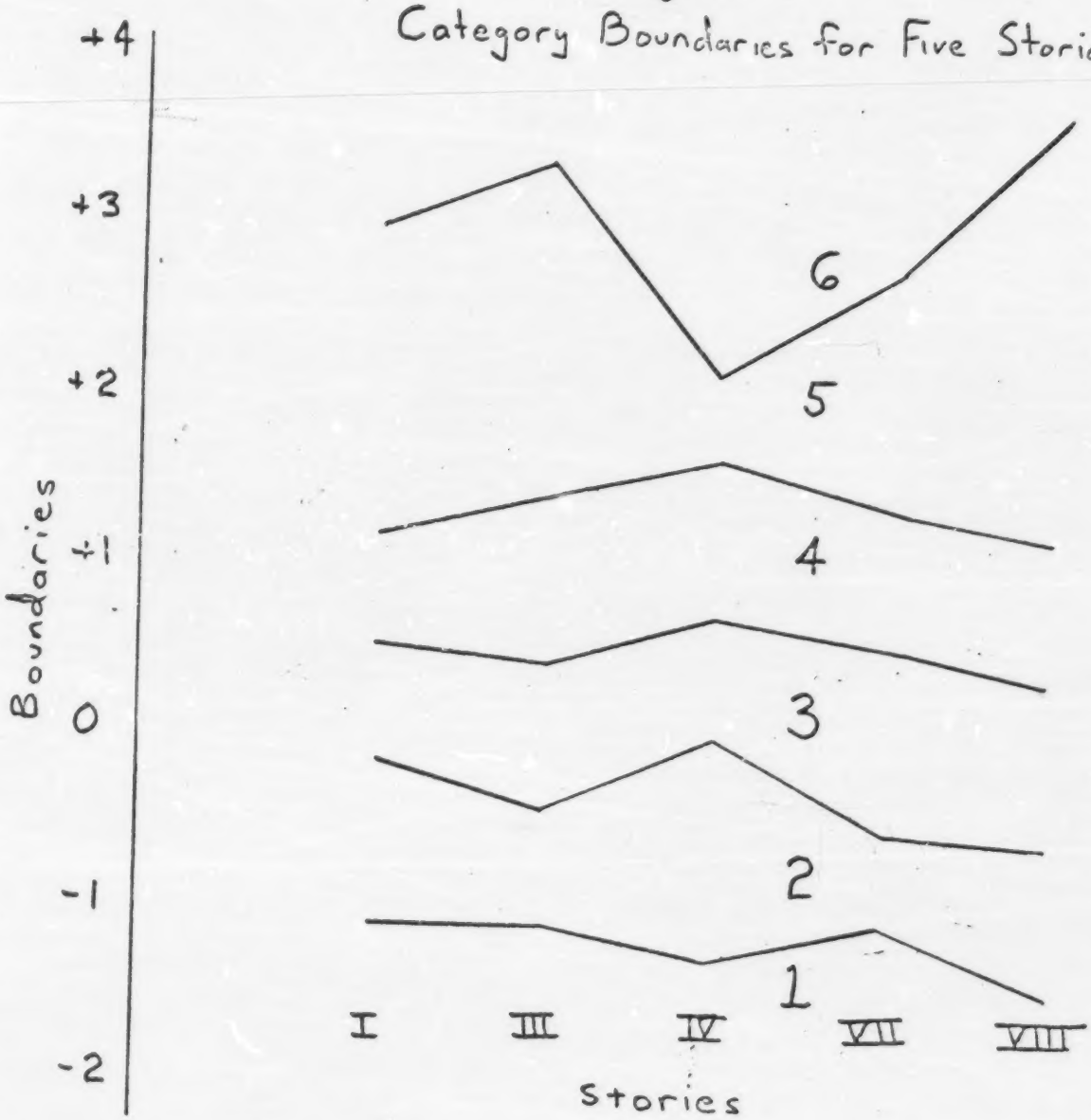


TABLE 3

EXAMPLES OF THE SAME GLOBAL SCORE
PRODUCING VARIOUS MORAL JUDGMENT LEVELS

Global Score	Pattern					M.J. Level
	I	II	III	VII	VIII	
	3	2	2	3	2	-1.2103
240	2	3	2	2	3	-0.8742
	2	2	2	3	3	-0.8694
	3	2	3	2	3	-0.6950
260	2	2	3	3	3	-0.5737
	3	2	2	3	3	-0.5616
280	2	3	3	3	3	-0.1612
	3	3	2	3	3	-0.1468

Figure 4

Probability that a subject with MJL, θ, will cross a boundary between stages

(Story VIII)

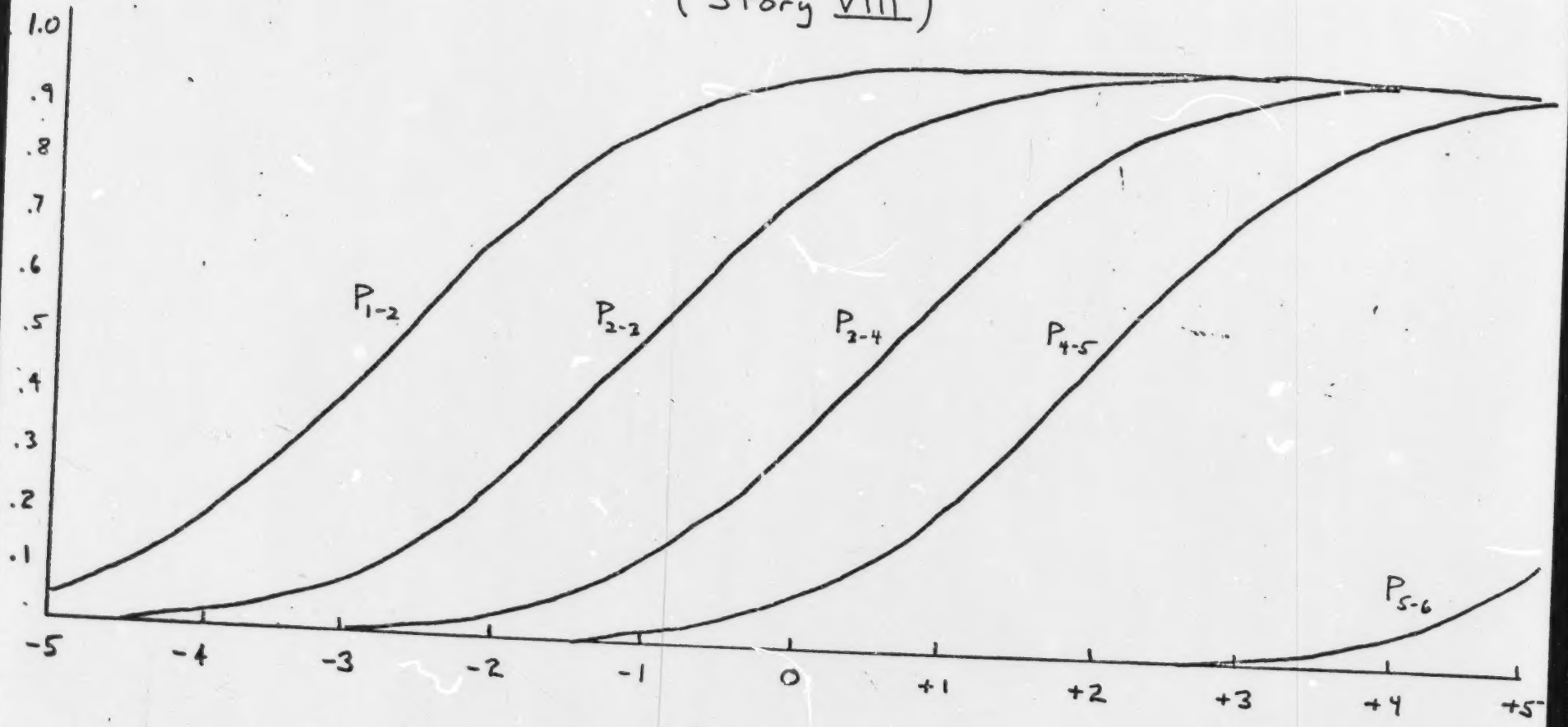


Figure 5

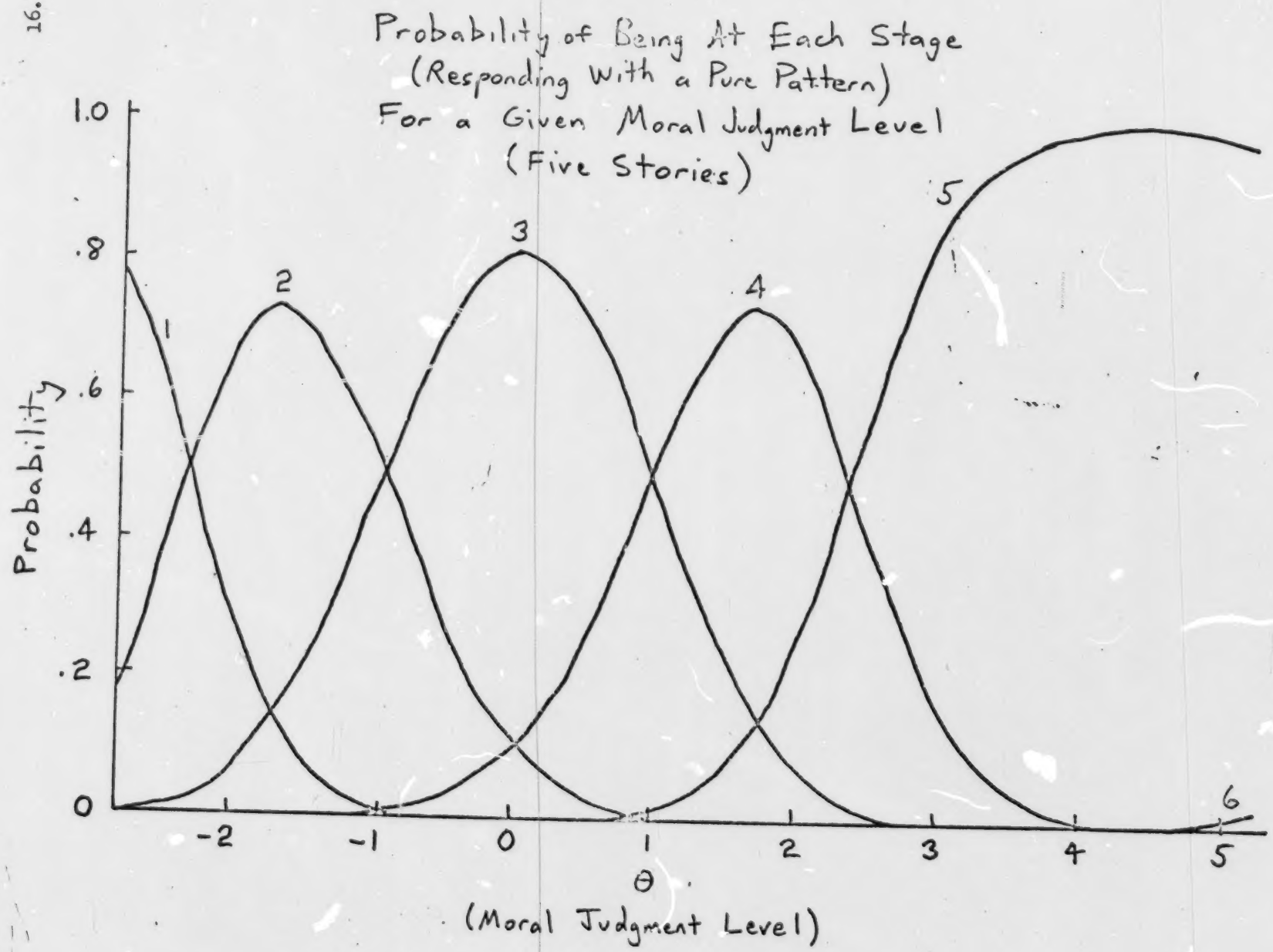

Probability of Being At Each Stage
(Responding With a Pure Pattern)
For a Given Moral Judgment Level
(Five Stories)
Probability
1.0
.8
.6
.4
.2
0
1
2
3
4
5
6
-2
-1
0
1
2
3
4
5
θ
(Moral Judgment Level)

Figure 6

The Issues

A The Legal System

A_{II} Rules of Law (and rules of custom with quasi-legal force) (norms)

G Relations of Punishment and of Legal Judgment (relations)

E Civil Rights, Basic Rights to Liberty and Equity Which Law Must Protect (value)

L Public Welfare and Order (value)

B The System of Conscience

B_{III} Obligation and Blame, Modes of Obligation, Responsibility Blame where rules and values conflict (norms)

B_{I} Guilt, fear, anxiety and moral character-maintenance as motivating choice (value)

B_{V} Personal Ethical Theory

C The Affectional System

C_{II} Role Stereotypes and norms of good family and friendship roles (norms)

C Relations of affection and concern for welfare (relations)

K Love, intimacy and sex (value)

D The Leadership and Power System

D_{II} Civil Stereotypes and norms of good authority and good citizen-follower roles (norms)

D Relations of Authority, Respect, and Leadership (relations)

E Civil Rights (value)

L Public Welfare and Order (value)

F The Economic System

F_{VI} Work Role Stereotypes and norms (norms)

F Relations of Contract and Equity (relations)

I Property (value)

J Truth and Trust (value)

H Life (value)

Table 4

Issue Parameters

Discriminating Powers

A	B_I	B_{III}	C	D	F	G	H	I
.885	.857	.920	.924	.854	.850	.915	.734	.892

Stage Boundaries

A	-1.69370	-.91291	-0.08095	1.40074	2.48643
B_I	-1.79661	-.24663	.41419	1.21009	2.09468
B_{III}	-1.94496	-1.01380	.34327	1.27971	2.33742
C	-2.00315	-.95035	.48764	1.33877	2.72390
D	-1.93982	-1.38537	-.03832	1.38537	5.47119
F	-2.33377	-.83232	.14799	1.18683	2.33377
G	-1.69370	-.84162	.06473	1.26338	2.72390
H	-2.01451	-.66587	.38087	1.40114	2.44087
I	-2.48064	-.75132	-.05345	1.53579	2.71852

Figure 7

Category Boundaries for Nine Issues

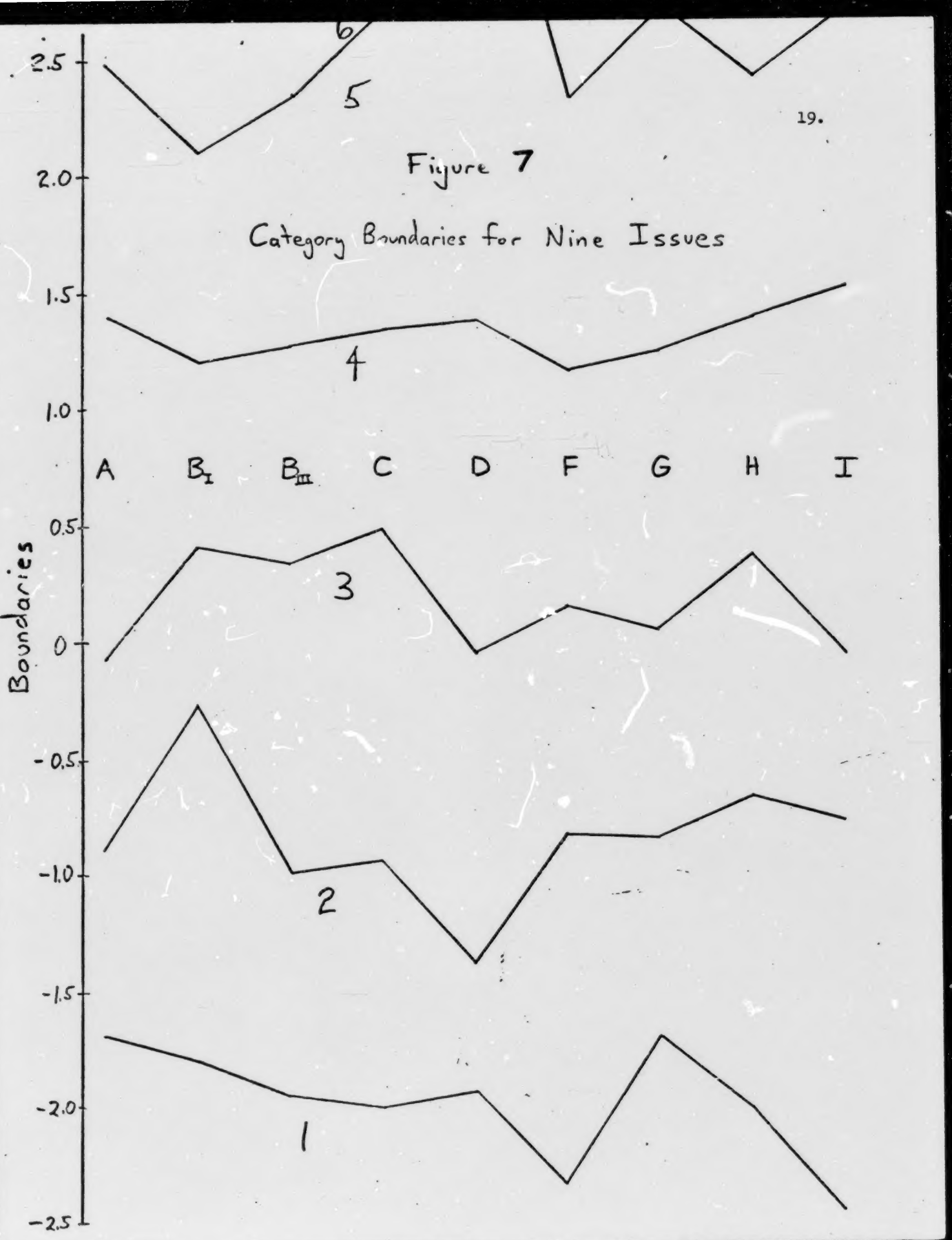


Figure 8

Probability that a subject with MJL, θ, will cross a boundary between stages

(Issue C)

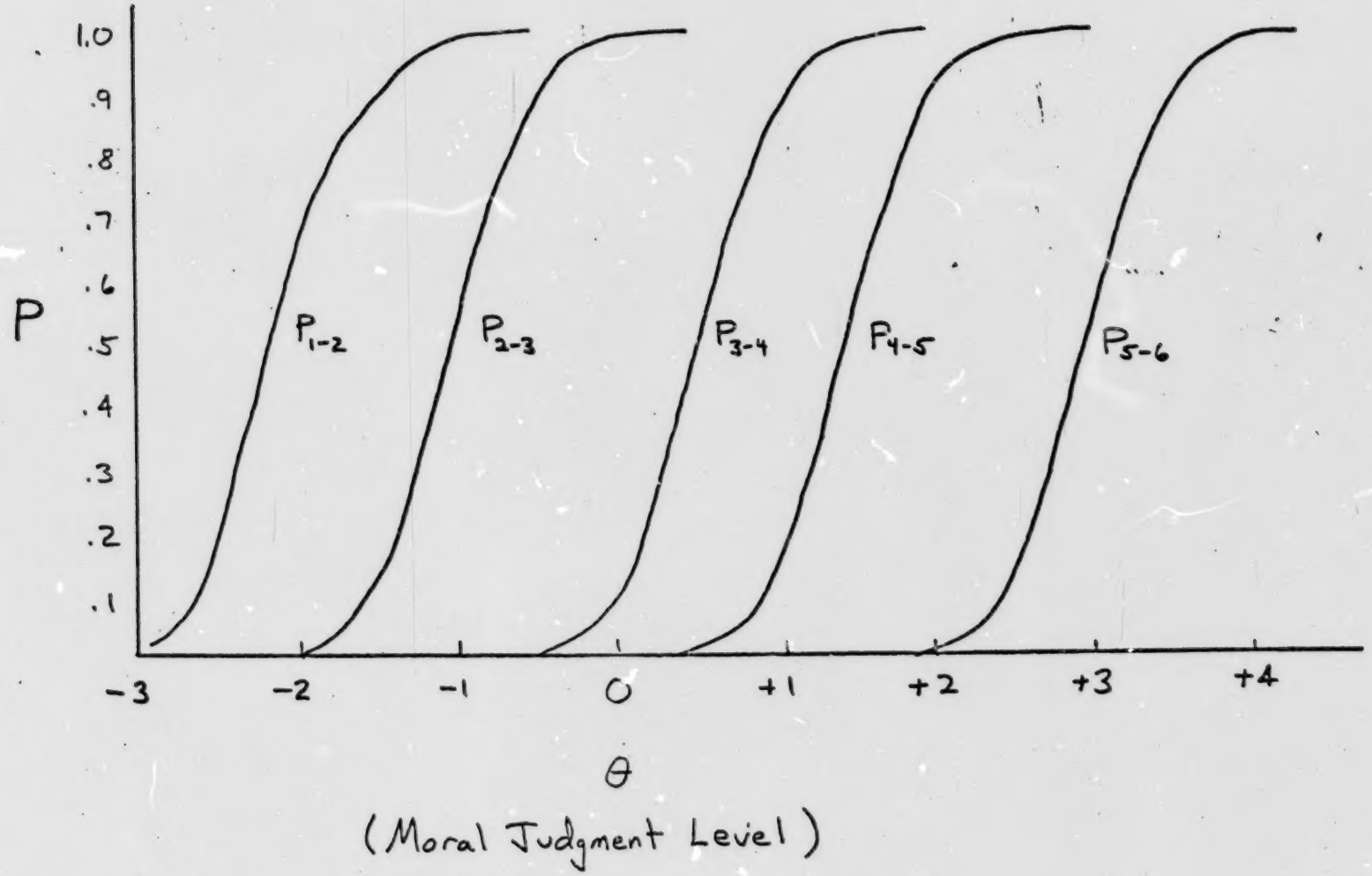
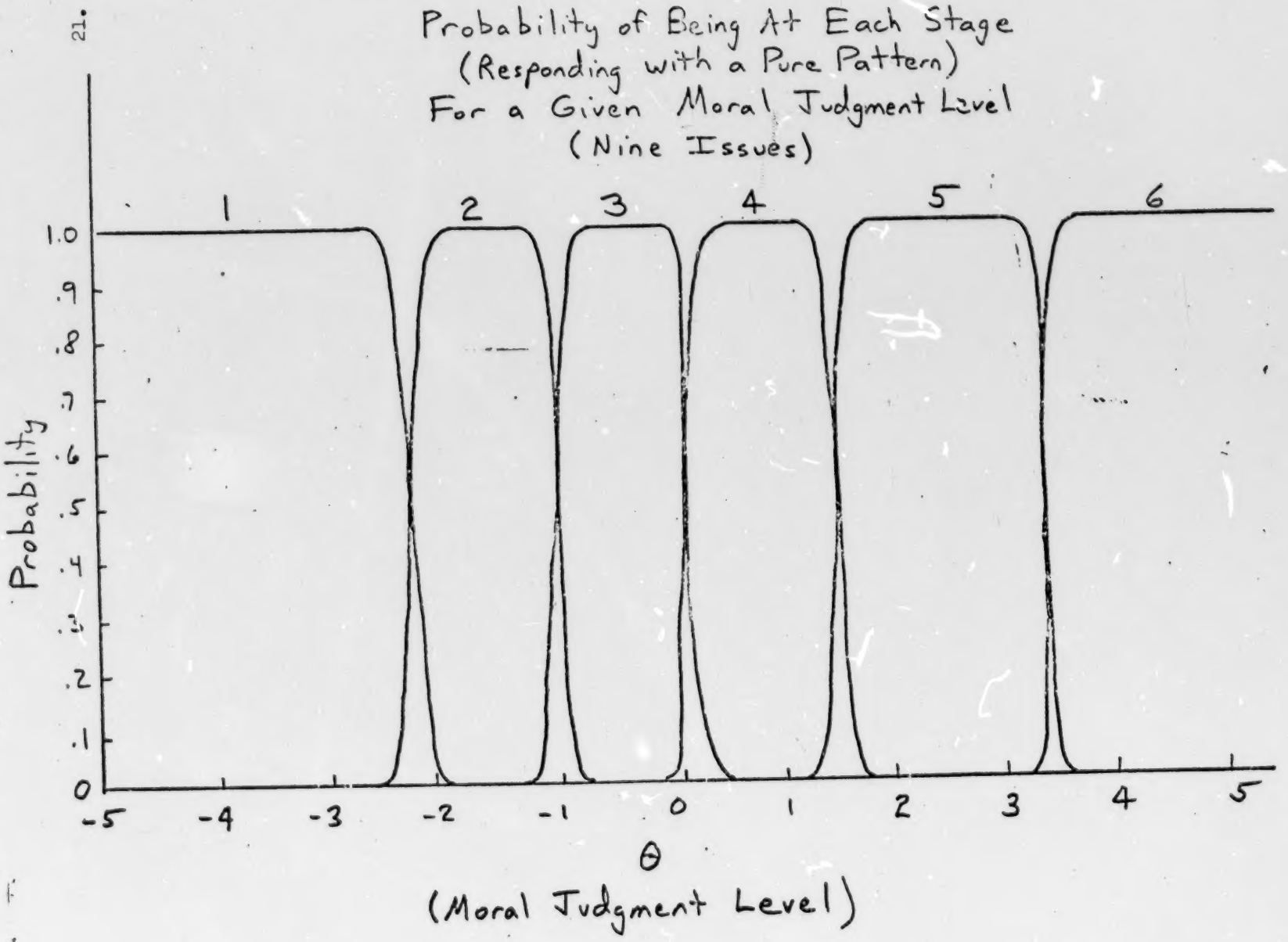


Figure 9


Probability of Being At Each Stage
(Responding with a Pure Pattern)
For a Given Moral Judgment Level
(Nine Issues)
1
2
3
4
5
6
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
Probability
-5
-4
-3
-2
-1
0
1
2
3
4
5
θ
(Moral Judgment Level)